The history and theory of the doubly labeled water technique[1,2]

John R Speakman

ABSTRACT Scientists have been measuring energy expenditure by using gas exchange for the past 200 y. This technique is based on earlier work in the 1660s. Gas exchange in respirometers provides accurate and repeatable measures of resting metabolic rate. However, it is impossible to duplicate in a respirometry chamber the diversity of human behaviors that influence energy expenditure. The doubly labeled water technique is an isotope-based method that measures the energy expenditure of unencumbered subjects from the divergence in enrichments of 2 isotopic labels in body water—1 of hydrogen and 1 of oxygen. The method was invented in the 1950s and applied to small animals only until the early 1980s, mostly because of the expense. Since 1982, when the first study in humans was published, its use has expanded enormously. Although there is some debate over the precise calculation protocols that should be used, the differences between alternative calculations result in relatively minor effects on total energy expenditure estimates (≈6%). Validation studies show that for groups of subjects the method works well, but that precision is still relatively poor (8–9%) and consequently the method is not yet sufficiently refined to provide estimates of individual energy expenditures. *Am J Clin Nutr* 1998; 68(suppl):932S–8S.

KEY WORDS Doubly labeled water, gas exchange, respirometry, energy expenditure, carbon dioxide production, indirect calorimetry, dilution spaces

INTRODUCTION

The knowledge that respiration and ventilation are essential functions of life dates back to at least biblical times. In the *Old Testament, Book of Psalms*, for example, it is stated with respect to animals, "When thou takest away their breath they die" (Psalm 104). Although it was clear to ancient humans that animals had to breathe to live, the exact physiologic function of breathing was obscure. For many years it was widely believed that inspired air served to cool the heat generated by a fire burning in the heart (1). Scientific study of animal respiration only commenced in the 1600s. In 1660 Robert Boyle observed that mice that had been sealed into bell jars expired at the same time that a burning flame became extinguished. Boyle thus established 2 important principles: the equivalence of fire and life as combustion processes and the requirement of air to support these processes.

Less well known, but probably of greater significance, was the work of John Mayrow. In 1668 Mayrow placed mice in sealed jars over water and charted the change in volume of the air inside the jar from the change in the movement of the water. He observed that mice died when they had consumed about one-fourteenth of the air in the bell jar. Mayrow accordingly established the idea that the air consists of different parts, only some of which are usable for the process of respiration. A significant discovery, for which he is not generally credited, is the invention of a chamber that allowed the quantification of the consumed portion; this was the first respirometer. The significance of Mayrow's observations was not widely recognized, mostly because of the prevailing alchemic framework within which they were interpreted.

A century after the innovative work of Boyle and Mayrow, the French chemists Lavoisier and Seguin started systematic investigations of respiration as a process analogous to combustion. In the intervening century, important advances had been made toward the understanding of the chemistry of gases: in 1757 Joseph Black had discovered carbon dioxide ("fixed air"), and in 1774 Joseph Priestly had discovered oxygen. The procedures used by Lavoisier and Seguin mimicked closely those developed by Mayrow, the key difference being the framework within which the observations were interpreted. The methods involved confining animals or humans in chambers to quantify their consumption of oxygen and production of carbon dioxide. The chambers they used were relatively unsophisticated. Animals were still confined in bell jars, and human subjects were confined in varnished silk bags secured around the nose and lips with pitch.

Lavoisier and Seguin made several important discoveries about oxygen consumption. First, they found that larger persons consume more oxygen than smaller persons. Second, they found that people sitting quietly at rest consumed less oxygen than those standing up or moving about. Finally, they discovered that after a meal, oxygen consumption was elevated even in a person sitting at rest. Perhaps most importantly, the work of Lavoisier and Seguin established the methodology of indirect calorimetry that has remained the benchmark method for the quantification of animal and human energy expenditure to this day.

Since the end of the 18th century, the chambers in which animals and humans are confined have become increasingly sophisticated. Moreover, sealed systems have been replaced with open flow systems linked to advanced gas analysis equipment, so

[1] From the Department of Zoology, University of Aberdeen, Scotland, United Kingdom.

[2] Reprints not available. Address correspondence to JR Speakman, Department of Zoology, University of Aberdeen, Aberdeen AB24 2TZ, Scotland, United Kingdom. E-mail: J.Speakman@abdn.ac.uk.

Am J Clin Nutr 1998;68(suppl):932S–8S. Printed in USA. © 1998 American Society for Clinical Nutrition

The American Journal of Clinical Nutrition

Downloaded from ajcn.nutrition.org by guest on January 31, 2013

subjects no longer have to die to provide a measurement. Nevertheless, it is clear that no matter how intricately designed such chambers become they will never be able to reproduce the complexity of activities in which people are engaged as they go about their routine lives. Yet, our understanding of many of the nutritional problems facing humans depends on reliable and accurate measurements of their energy demands as they go about their everyday business. The inadequacy of traditional calorimetry apparatus has been recognized for a long time, and there have been many attempts to develop methods, such as heart-rate monitoring, that enable the energy demands associated with free-living activities to be monitored (2–4). The doubly labeled water (DLW) technique is a method that allows the measurement of energy demands of free-living subjects. The success of this method prompted Prentice (5) to remark that its development was as significant an event in the history of animal and human nutrition as the work of Lavoiser and Seguin had been.

HISTORY OF THE DOUBLY LABELED WATER TECHNIQUE

The discovery of rare heavy isotopes of oxygen and hydrogen in the late 1920s and the 1930s transformed many areas of biology. This is because these substances were chemically and physically almost identical to the common light forms of the elements. The rare isotopes were thus ideal as tracers for the behavior of oxygen and hydrogen, and therefore water, in the body. Until their discovery, scientists relied on water-soluble dyes to investigate water dynamics. For example, known amounts of the dye Evans blue were injected into subjects to see how diluted it became and thus evaluate their plasma blood volume (eg, 6). In the early years after the discovery of isotopes, the major problems with their use related to their expense, lack of availability, and, for stable forms, the lack of equipment capable of accurately measuring their abundance. These problems were solved principally as a byproduct of the need to understand more about the behavior of isotopes and to artificially enrich them during the development of atomic weapons at the end of the Second World War.

By 1949, Lifson et al (7), at the University of Minnesota, had performed several experiments in mice that involved injecting them with stable isotopes of oxygen in water or forcing them to breathe air enriched with ^{18}O. These experiments clearly showed that the oxygen in body water was in complete isotopic equilibration with the oxygen in respiratory carbon dioxide. This equilibration occurred principally because of isotope-exchange reactions catalyzed by carbonic anhydrase during blood gas transport. The consequence of these exchange reactions is that a dose of oxygen isotope introduced into the body of an animal is removed by the flow rates of water, inspired oxygen, and expired carbon dioxide through that body. Although this consequence was not stated by Lifson et al (7), its significance was clearly not lost on the authors. If an isotopic label of oxygen is eliminated from the body by both carbon dioxide and water, an isotopic label of hydrogen would be eliminated only by the water. Thus, the difference in the elimination of the 2 labels simultaneously introduced into the animal (thus DLW) would provide a measure of the carbon dioxide production and hence indirectly the energy expenditure.

It took 6 y to develop this simple theory into a working method. The seminal publication on the DLW technique, which involved comparisons of estimates of carbon dioxide production by standard respirometry and the DLW method in 15 laboratory mice, was published in 1955 (8). The technique provided an estimate that differed from the standard method by ≈2–3% on average, depending on the assumptions made about the behavior of isotopes both in the body and during elimination from it. It was immediately clear that the method worked and that it had a tremendous advantage over standard respirometry measurements. As Lifson et al (8) put it in the introduction to their 1955 paper, "a measurement is made possible merely by taking 2 blood samples to reconstruct the isotope elimination curves." This meant the animal could perform a whole variety of natural behaviors between the taking of the 2 samples unconstrained by confinement in a respirometry chamber. Yet it would be another 9 y before this potential advantage of the method was actually exploited to study metabolism of a wild animal (9). The method was used to measure the energy costs of flying in homing pigeons (*Columba livia*) (9) in 1964. The author was also based at Minnesota and associated with Lifson's group. Indeed, Lifson's group was the only group in the world to use the method between its original development and 1970, and during that time they produced only 9 papers using it, of which 6 were validations (all on small rodents), 2 were applications, and 1 was a complete theoretical analysis of the method and its underlying assumptions (10). This theoretical paper has formed the basis of all the subsequent theoretical discussions of the method. Few of the advances made since the mid-1960s are not addressed in this paper. By the mid-1990s, 30 y after publication, records from the *Science Citation Index* (Institute for Scientific Information, University of Aukland, New Zealand) indicate that this paper was cited at least 50 times/y.

In the 1970s, several other groups started to use the method. Studies by these groups exploited the ability of the method to measure the free-living energy demands of small animals. The first measures on a free-ranging mammal (pocket mouse, *Perognathus formosu*) were made in 1970 (11) and the first on a free-living reptile (lizard, *Sauromalus obesus*) in 1972 (12). All the animals on which the method was used between 1955 and 1975 weighed <1 kg. There were no attempts to apply the method to humans to evaluate free-living energy expenditures for 2 reasons. The first reason is that it was prohibitively expensive. In 1973, for example, it was estimated that to use the method on a single 70-kg human would cost ≈US$50 000 (13). For comparison, the most expensive production automobile in 1973 was the Rolls Royce Phantom VI that cost £17 817 (≈US$30 000 at current exchange rates). The second reason was that clinical problems involving disorders of energy balance (eg, obesity) were less prevalent in Western populations at the time (14, 15), and links between such disorders and other clinical problems were not fully understood. Thus, in the early 1970s we had available a method that would have cost the equivalent of 2 Rolls Royce automobiles per subject to implement and no immediately apparent clinical problem of sufficient importance to which the results could be applied.

The cost of ^{18}O declined throughout the 1970s, and developments in mass spectrometer technology made it feasible to label subjects at much lower levels and still retain precision and accuracy of analysis. By 1980, the costs of a study of a 70-kg subject had declined to ≈US$2000. Moreover, the impetus to perform such studies was increasing because the prevalence of morbid obesity increased during the previous decade, and links between

The American Journal of Clinical Nutrition

Downloaded from ajcn.nutrition.org by guest on January 31, 2013

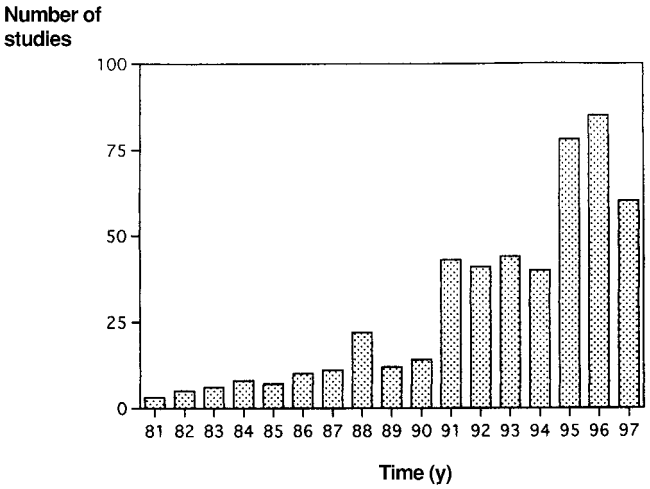

FIGURE 1. Number of studies in peer-reviewed journals (excluding abstracts) that used the doubly labeled water technique in the years 1981–1997 (through June) from the *Science Citation Index* (Institute for Scientific Information, University of Aukland, New Zealand). Since the first study in humans in 1982 the use of the technique has continued to grow.

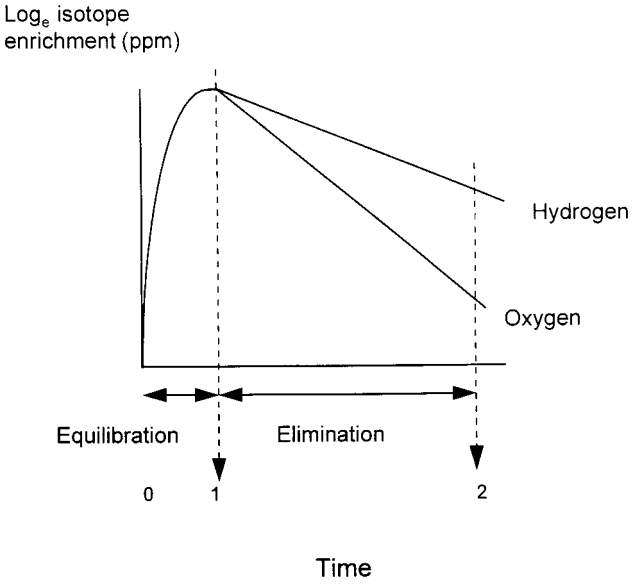

FIGURE 2. Theoretical time course of enrichments of isotopes of oxygen and hydrogen in body water after administration at time zero. Over an equilibration period of several hours the isotope enrichments rise to a peak. If the amount of administered isotope is known, the peak enrichment can be used to establish the volume of dilutant material in the body (called the dilution space) by using the dilution principle. After equilibration, the isotopes are washed out of the body along an exponential track (which is linear when expressed as the log of the enrichment above background as shown here). The oxygen isotope leaves the body faster than the hydrogen isotope because it is washed out of the body by water and carbon dioxide. A measure of carbon dioxide production and hence energy expenditure is made from the divergence of the enrichments between time 1 and time 2. During this interval (4–14 d in a typical study in adult humans) the subject is free to engage in usual activities without the need to be confined in a respirometery chamber. Note: the relative timings of the equilibration and elimination phases are not drawn to scale and the extent of divergence is exaggerated for illustrative purposes.

obesity, heart disease, and other clinical problems were becoming increasingly apparent (eg, 16, 17). It seems plausible that increasing financial feasibility combined with practical applicability led to the first use of the method in humans. Indeed, previous reviews of the history of the method have painted this picture (eg, 5). The facts, however, do not accord with this interpretation.

The method was apparently independently discovered by accident in the early 1980s because of an anomaly that occurred during attempts to crosscheck hydrogen turnover estimates of water requirements in humans (18). When the first DLW measurement was made on a man in 1980, Schoeller had no notion that he was using a method that had been developed 25 y previously and by that time was in relatively widespread use by researchers studying the energy demands of small animals. The first validation study using the DLW method in humans was published in 1982 (19). Validation studies of the method in humans proliferated in the literature over the next 6 y. These validations included studies of infants (20, 21) and adults (22, 23). The first application of the method in a group of free-living humans was published in 1985 and addressed the energy expenditure of obese and lean subjects living in Cambridge, United Kingdom (24).

Since the mid 1980s, use of the method has expanded enormously (**Figure 1**) in concert with a proliferation of validation work (in both humans and animals). By the mid-1990s the technique was being used in ≈70–90 publications/y. Attention has focused on the underlying assumptions on which the method is based, and refinements in the calculation method have been proposed (eg, 25) and widely adopted. The method has been applied to an ever increasing array of wild animals of an expanding range of body sizes. As of June 1997, the largest species to which the method was applied was the polar bear (*Ursus maritimus*), weighing 250 kg (26), and the smallest was the honeybee (*Apis mellifera*), weighing only 83 mg (TJ Wolf, unpublished observations, 1996). In addition, the method has been applied to measure the energy costs associated with several clinical conditions (paraplegia, parenteral nutrition, and burns); routine activities of a range of human subjects of varying body masses, ages, and activities; and more bizarre human endeavors, such as climbing Mount Everest (27) and competitive bicycle racing for 21 d (28). A review of the measurements made on free-living human subjects was published in 1996 (29).

THEORY OF THE DOUBLY LABELED WATER METHOD

The fundamental basis of the DLW method is that oxygen turnover in a body is dominated by the flow of water through the body as well as inspired oxygen and expired carbon dioxide. The turnover of body hydrogen, however, is dominated only by the flow of water through the body. Consequently, the difference between the turnovers of oxygen and hydrogen provides a measure of the excess efflux of oxygen that is equivalent to the production of carbon dioxide (**Figure 2**).

In practice, the method involves introducing isotopes of both oxygen and hydrogen into the body water to quantify the turnover of the body pools of the 2 elements. The size of the pool in which the isotopes are distributed and the rates of their elimination provide direct estimates of the flow of compounds carrying the 2 labels out of the body. The rates of isotope elim-

The American Journal of Clinical Nutrition

Downloaded from ajcn.nutrition.org by guest on January 31, 2013

ination are quantified from the exponential decline in isotope enrichment in the body water (Figure 2). Given the isotope enrichment in a sample taken shortly after isotope administration (E_I) and a second sample taken some time later (E_f) the elimination rate (k) is calculated from the extent that these 2 body enrichments exceed the background isotope level (E_b). Generally, for any isotopic label,

$$k = [\mathrm{Log}_e(E_I - E_b) - \mathrm{Log}_e(E_f - E_b)]/t \quad (1)$$

In its very simplest form, therefore, the DLW method suggests that the elimination rates of oxygen (k_o) and hydrogen (k_d) turning over in the body water (N) are related to the rate of carbon dioxide production (rCO_2) by the formula

$$rCO_2 = (N/2) \cdot (k_o - k_d) \quad (2)$$

where the value 2 is a constant reflecting the fact that 1 mol CO_2 removes 2 atoms of O_2 and 1 mol H_2O leaving the body removes only 1.

Fractionation

Unfortunately, rare isotopes do not behave exactly like their common counterparts. In particular, when compounds react or when they are vaporized, the lighter species tend to require less activation energy and thus react or vaporize more readily than the heavier species. This discriminatory effect is termed fractionation, and a constant value, called the fractionation factor, expresses the extent of the discrimination in any particular process. The fractionation factor represents the ratio of the enrichment of the rare isotope in the 2 phases. For example, the fractionation factor for hydrogen in gaseous water compared with liquid water at 25°C is 0.94, meaning the enrichment of the rare isotope (deuterium) in the gaseous phase is only 94% of that in the liquid. The existence of fractionation complicates the simple model because fractionated material leaving the body does not take with it the same amount of label as that left behind in the body. The effect is further complicated by the fact there are 2 different modes of fractionation (kinetic and equilibrium) and they discriminate against the heavy species to different extents.

The importance of fractionation was recognized by Lifson et al (8), who generated a correction based on some simplifying assumptions about the nature of the fractionating process. They assumed that the carbon dioxide leaving the body was all fractionated and that half the water losses were also fractionated. They then used the equilibrium fractionation factors at 25°C measured in vitro for all the fractionating processes to derive the following equation:

$$rCO_2 = (N/2.08) \cdot (k_o - k_d) - 0.015 \cdot k_d \cdot N \quad (3)$$

In validation, this equation performed much better than the equation that ignores the fractionating effects (*Eq 2* above) (8). However, it is clear that the corrections are based on several dubious assumptions. Schoeller et al (25) overhauled the fractionation correction, making more realistic assumptions for the processes involved in humans. Specifically, they used the fractionation factors pertinent at 37°C rather than 25°C and assumed that fractionated water losses would occur in proportion to the rate of carbon dioxide production. The resultant equation was

$$rCO_2 = [(N/2.076) \cdot (k_o - k_d)] - 0.0246 \cdot 1.05 \cdot N(k_o - k_d) \quad (4)$$

Pool size models

In the original formulation of the model, Lifson et al (8) assumed that the oxygen and hydrogen in the body were turning over in the body water (N). However, both oxygen and hydrogen exist in other pools in the body, and the sizes of these extra exchanging pools are not equal. Generally, the hydrogen pool exceeds the oxygen pool by 3–5%, which in turn exceeds the body water pool by ≈1%. Kinetic analyses indicate that the true estimate of the flow of materials is the product of the elimination rate and the size of the pool in which the isotopes are turning over. Thus, rather than multiplying the difference in elimination rates by the body water (N), it was thought that each elimination constant should perhaps be multiplied by its own dilution space in the body, N_o and N_d respectively. Therefore, Schoeller et al (25) modified equation 4 to

$$rCO_2 = (1/2.076) \cdot (N_o \cdot k_o - N_d \cdot k_d) - 0.0246 \cdot 1.05 \cdot N(k_o - k_d) \quad (5)$$

In this equation, the first term includes the elimination constants each multiplied by the respective dilution spaces, but in the second term they are still multiplied by the body water (N). Schoeller (30) modified the equation so that in both terms the elimination constants are multiplied by their respective dilution spaces, as follows:

$$rCO_2 = (1/2.076) \cdot (N_o \cdot k_o - N_d \cdot k_d) - 0.0246 \cdot 1.05 \cdot (N_o k_o - N_d k_d) \quad (6)$$

However, the justification for this modification was questioned recently (31) on the basis that the value 1.05 in equation *5* was included to account for the difference between $N(k_o - k_d)$ and $(N_o k_o - N_d k_d)$, and consequently, by including 1.05 and $(N_o k_o - N_d k_d)$, a double correction for the pool size difference is made. However, the practical consequences of these differences for a typical study has been shown to be relatively trivial (<0.5%) (31).

Lifson et al (8) evaluated the effects of using both pool sizes rather than a single pool size in their validations of mice and concluded that it did not improve the fit of DLW estimates to those made by indirect calorimetry. However, this may have been because the effects of the different pool sizes on the calculation were already being accounted for by the erroneous assumptions concerning fractionation. Nevertheless, recalculation of all the original validation studies from the 1950s using the more modern equations confirmed that, in small rodents, the equation that includes only a single pool performs better than the 2-pool model (31).

Coward et al (32) and Schoeller et al (25) reevaluated the effects of using different pool models in humans and came to the conclusion that using the 2-pool model provided a better fit to calorimetry data than did using the single-pool model. This difference between the appropriateness of the different models in different-sized subjects appears to reflect a real biological difference between small rodents and humans.

In fact, in addition to a small extra exchangeable pool, there is also a small extra hydrogen elimination route. The excess hydrogen space in adult humans is similar to that in adult small mammals (being 3–4% greater than the oxygen space in both) but the

The American Journal of Clinical Nutrition

Downloaded from ajcn.nutrition.org by guest on January 31, 2013

extra elimination route is not; proportionally, it is much larger in the small mammal (33). Theoretical analyses of the pool structure show that the extent of the extra elimination offsets the effect of the extra pool (8, 33, 34). Thus, in small mammals with a high extra elimination route, the single-pool model works best, but in larger subjects the 2-pool model works best. Subsequent validations in animals of a range of body sizes confirmed this effect and suggest that the single-pool model is probably appropriate in animals weighing ≤5 kg (35).

The effect of body size on the appropriateness of different pool models raises the interesting question of what might be the most appropriate method to calculate the metabolism of small infants (<3–4 kg). In fact, validation studies in infants clearly show that the 2-pool formulation gives a close fit to the calorimetric data even in preterm infants weighing <2 kg (20, 21), although calculations using the single-pool model have not been assessed. It seems likely that the 2-pool model works in small humans because of a combination of 2 factors. First, small babies do not necessarily behave physiologically like adult small animals of the same body mass with respect to the magnitude of the subsidiary hydrogen elimination. Second, the extent of their extra hydrogen space is much smaller than that found in adults; consequently, the difference between the single- and 2-pool model is reduced. Indeed, in very small infants the ratio of hydrogen to oxygen dilution space does not differ from unity, thus 2-pool and single-pool models give identical results.

Using individual or population-based dilution spaces in the 2-pool model.

An issue that has emerged over the past decade is whether to use individual values of N_o and N_d in the 2-pool model equation or to replace these with a population-based average (25). If observed variation in N_o and N_d is biological, then it would make sense to use individual values for the dilution spaces. Alternatively, if the variation is mostly analytic it would make sense to use a population-based ratio of the 2 pools. Some authorities continue to use individual ratios. However, it is now widely accepted that the dominant source of variation in the observed pool size ratio (N_d/N_o) is analytic. This being the case, a population-based average would be better used in the equation. Schoeller et al (25), on the basis of a limited sample of animal studies, suggested that the oxygen pool exceeds N by 1.01 and the hydrogen pool exceeds N by 1.04, and thus the hydrogen pool exceeds the oxygen pool by 1.03. Therefore, the 2-pool model equation (6) can be reformulated as

$$rCO_2 = (N/2.076) \cdot (1.01k_o - 1.04k_d) - 0.0246 \cdot 1.05 \cdot N(1.01k_o - 1.04k_d) \quad (7)$$

where $n = (N_o/1.01 + N_d/1.04)/2$.

The appropriateness of the population-based average of 1.03 is debated. This value was rounded from a true value of 1.034 (25), but it was based on the relatively small sample size available at the time. Subsequent reevaluations using larger data sets have resulted in different population averages, ie, 1.0427 (36), 1.027 (37), 1.034 (38), and 1.035 (39). The differences in these estimates depend on the inclusiveness of the reviews. Some authors have concluded that analyses of deuterium directly from urine are inaccurate and have eliminated these from their compilations of data (eg, 38). The consequent differences in the formulation of the equation can result in biologically significant differences in the final estimate of carbon dioxide production (36, 38, 39).

I reviewed previously all the data available as of July 1996 on both animals and humans for this ratio (31). This survey revealed that across 15 studies in mammals, the mean (±SD) dilution space ratio was 1.041 ± 0.025 and across 9 studies in birds the ratio was 1.0427 ± 0.033. Across 19 studies comprising 590 adult humans the weighted mean was 1.0387 (31). This value includes all data, irrespective of whether analyses were performed directly on urine or on water distilled from the urine. Debate over the most appropriate population estimate is certain to evolve as more data are published. However, a consensus is likely to fall between the limits of 1.03 (25) and 1.0427 (36).

Validation

A summary of all the validation comparisons of DLW to indirect calorimetry in human subjects between 1982 and 1996 is provided in **Table 1**. It is not possible to compare the different formulations of the equations directly on a matched sample of subjects because the data necessary to recalculate early validations were not published in the papers. Overall (**Table 2**), across 201 subjects the Schoeller et al (25) equation, with the lowest population ratio (1.03), overestimated the simultaneous indirect calorimetry by 2.1%. In contrast, in 54 subjects the Speakman et al (36) equation, with the highest ratio (1.0427), underestimated simultaneous indirect calorimetry by 2.2%. The equation of Racette et al (38) with an intermediate ratio (1.034) overestimated by on average 1.2% (n = 36). A limited matched sample can be compared. In the same 36 individuals reviewed by Racette et al (38), the Schoeller et al (25) equation overestimated indirect calorimetry results by 3.3% and the Speakman et al (36) equation underestimated these results by 3.1%. Differences between the equations are consequently relatively minor, but should perhaps be borne in mind when relatively small differences are compared between studies that might have used different methods, or when comparisons are made between estimates of energy expenditure using DLW and other methods for determining energy demands, such as food intake.

It is important to recognize that the average comparisons across groups indicate only the efficacy of the method for defining the energy demands of groups. Individual comparisons of DLW to indirect calorimetry have much greater deviations than the group means. Across studies the precision is ≈10% on average (Table 1). Thus, at the extremes, some individuals may have DLW estimates >20% divergent from simultaneous indirect calorimetry. There is some evidence that part of this variation can be attributed to analytic problems in certain laboratories; a multiple laboratory comparison (40) showed wide discrepancies in the isotope enrichment determinations of the same samples analyzed by different groups. Nevertheless, precision in the best laboratories is still relatively poor and the method is not yet sufficiently refined to make confident estimates of individual energy requirements. Repeatability of measurements made by the method at ≈6–10% confirm this (53).

The reasons for this lack of precision are not yet entirely clear. Precision may improve as the technical instrumentation for mass spectrometric determination of isotope abundances improves. However, in many circumstances the extents of deviations of estimates made by DLW and simultaneous indirect calorimetry already exceed the theoretical precision of the DLW measurement, given the precisions of the component analyses (54). This

The American Journal of Clinical Nutrition

Downloaded from ajcn.nutrition.org by guest on January 31, 2013

TABLE 1
Validation studies of the doubly labeled water method in human subjects in comparison to indirect calorimetry (IC) and food intake/mass balance (FI) or parenteral nutrition (PN)[1]

n	*n*obs	Mass	Com	Eqn	Abs	Prec	Arith	*n*+	*n*−	Min	Max	Reference
		kg			*%*		*%*					
4	4	78.1	FI	L+M	5.0	5.6	+2.1	3	1	−5.8	7.1	19
2	2	—	IC	?L+M	3.5	3.5	−2.5	1	1	−6.0	1.0	40
1	1	70.5	IC	L+M	—	—	−4.6	0	1	—	—	23
5	5	72.9	IC	L+M	8.5	6.8	+5.9	4	1	−6.5	14.1	22
4	4	—	IC	C	—	2.0	+2.0	—	—	—	—	31
6	6	68.9	IC	L+M	8.0	9.0	+5.0	5	1	—	—	25
14	14	—	IC	S	—	7.0	+1.0	10	4	—	—	25*
14	14	—	IC	L+M	—	8.0	+4.0	11	3	—	—	25*
5	5	53.2	PN	S	5.7	5.9	+3.3	4	1	—	—	41
4	4	1.59	IC	C	4.2	4.8	−1.4	1	3	−4.8	5.8	20
9	10	3.2	IC	S	—	6.2	−0.9	5	5	—	—	21
5	5	—	IC	S	4.3	3.9	+1.4	3	2	—	—	42
4	8	—	IC	S	4.8	7.0	−1.0	3	5	—	—	43
8	8	2.67	IC	S	10.7	12.9	−8.7	—	—	−29.3	4.8	44
33	33	—	ICFI	L+M	6.9	6.6	+4.3	26	7	−8.7	20.2	30*
33	33	—	ICFI	S	5.1	6.3	+0.6	21	12	−12.1	16.7	30*
1	1	66.9	FI	S?	—	—	−0.8	0	1	—	—	45
4	4	84.6	IC	S	2.5	2.5	+15.3	4	0	—	—	46
12	12	108.0	IC	S	6.4	6.9	−2.5	5	7	—	—	47
27	27	79.8	FI	S	—	0.7	+2.5	17	10	—	—	48
8	8	74.1	FI	S	9.7	14.4	+2.2	5	3	−35.0	9.0	49
9	9	71.8	IC	S	—	4.6	+1.8	6	3	−4.0	12.0	50
6	6	50.1	IC	C	15.0	16.8	+7.0	4	2	−23.0	32.0	51
18	18	—	IC	L+M	—	—	+23.1	—	—	—	—	52
18	18	—	IC	S	—	10.7	+10.3	—	—	—	—	36*
18	18	—	IC	SNG	—	9.3	−0.4	—	—	—	—	36*
36	36	—	IC	S	—	8.9	+3.3	—	—	—	—	37*
36	36	—	IC	SNG	—	7.9	−3.1	—	—	—	—	37*
36	36	—	IC	R	—	8.5	+1.2	—	—	—	—	37*

[1]Abs, absolute mean deviation (average percentage error with all signs made positive); Arith, arithmetic percentage error including signs; C, Coward 2-pool model using individual dilution spaces; Eqn, equation used to calculate the doubly labeled water results; L+M, Lifson and McClintock single-pool model; Mass, mean mass of subjects in kg; Max, maximum deviation; Min, minimum deviation; *n*+, number of observations in which the DLW method exceeded the reference method; *n*obs, number of comparisons; *n* +, number of observations in which the DLW method exceeded the reference method; *n*−, number of observations in which the DLW method was lower than the reference method; Prec, precision calculated as the SD of individual estimates used to derive Arith; R, Racette et al, 1994 (2-pool model fixed ratio 1.034); S, Schoeller et al, 1986, equation *A6* (2-pool model fixed pool ratio 1.03); SNG, Speakman et al, 1993 (2-pool fixed pool ratio 1.0432); *, review of previously published validations.

suggests that analytic considerations alone do not underpin all of the observed variability. The ultimate precision of the method may then depend on individual variability in competing physiologic reactions involving the isotopic labels, dilution spaces, background drift, and fractionation effects, about which we now know relatively little.

TABLE 2
Summary statistics of data in Table 1[1]

Equation	Mean arith	Prec	*n*
L+M	+7.7	7.4	87
S	+1.8	8.2	201
SNG	−2.2	9.3	54
R	+1.2	8.5	36

[1]Mean arith, weighted mean arithmetic error across all validation studies; Prec, weighted mean precision; *n*, number of individual comparisons included in the calculation; L+M, Lifson and McClintock single-pool model; R, Racette et al, 1994, (2-pool model fixed ratio = 1.034); S, Schoeller et al, 1986, equation *A6* (2-pool model fixed pool ratio 1.03); SNG, Speakman et al, 1993 (2-pool fixed pool ratio 1.0432).

REFERENCES

1. Brown JC. A history of chemistry. 2nd ed. London: Churchill, 1920.
2. Livingstone MBE, Prentice AM, Coward WA, et al. Simultaneous measurement of free-living energy expenditure by the doubly labeled water method and heart-rate monitoring. Am J Clin Nutr 1990;52:59–65.
3. Livingstone MB, Coward WA, Prentice AM, et al. Daily energy expenditure in free-living children: comparison of heart-rate monitoring with the doubly labeled water ($^2H_2{}^{18}O$) method. Am J Clin Nutr 1992;56:343–52.
4. Schulz S, Westerterp KR, Bruck K. Comparison of energy expenditure by the doubly labeled water technique with energy intake, heart rate, and activity recording in man. Am J Clin Nutr 1989; 49:1146–54.
5. Prentice AM. The doubly labelled water method for measuring energy expenditure, technical recommendations for use in humans. Nahres-4. Vienna: International Atomic Energy Agency, 1990.
6. Gibson JG, Evelyn KA. Clinical studies of the blood volume IV. Adaptation of the method to the photoelectric microcolorimeter. J Clin Invest 1938;17:153–9.
7. Lifson N, Gordon GB, Visscher MB, Nier AO. The fate of utilised molecular oxygen and the source of the oxygen of respiratory carbon

The American Journal of Clinical Nutrition

Downloaded from ajcn.nutrition.org by guest on January 31, 2013

dioxide, studied with the aid of heavy oxygen. J Biol Chem 1949;180:803–11.
8. Lifson N, Gordon GB, McClintock R. Measurement of total carbon dioxide production by means of $D_2{}^{18}O$. J Appl Physiol 1955;7: 704–10.
9. LeFebvre EA. The use of $D_2{}^{18}O$ for measuring energy metabolism in *Columba livia* at rest and in flight. AUK 1964;81:403–16.
10. Lifson N, McClintock R. Theory of use of the turnover rates of body water for measuring energy and material balance. J Theor Biol 1966;12:46–74.
11. Mullen RK. Respiratory metabolism and body water turnover rates of *Perognatus formosus* in its natural environment. Comp Biochem Physiol 1970;32:259–65.
12. Nagy KA. Water and electrolyte budgets of a free-living desert lizard, *Sauromalus obesus*. J Comp Physiol 1972;79:39–62.
13. Mullen RK. The $D_2{}^{18}O$ method of measuring the energy metabolism of free-living animals. In: Gessaman JA, ed. Ecological energetics of homeotherms. Logan, UT: Utah State University Press, 1973; 32–43.
14. Kuczmarski RJ, Flegal KM, Campbell SM, Johnson CL. Increasing prevalence of overweight among US adults—the National Health and Nutrition Examination Surveys 1960–1991. JAMA 1994; 272:205–11.
15. Prentice AM, Jebb SA. Obesity in Britain—gluttony or sloth? Br Med J 1995;311:437–9.
16. Keys A. Overweight, obesity, coronary heart disease, and mortality. Nutr Rev 1980;38:297–307.
17. Garrow JS. Treat obesity seriously: a clinical manual. Edinburgh: Churchill Livingstone, 1981.
18. Schoeller DA. Development of the doubly labelled water method for measuring energy expenditure. Proceedings of the AO Nier Symposium on Inorganic Mass Spectrometry. Durango Colorado: 1993.
19. Schoeller DA, van Santen E. Measurement of energy expenditure in humans by doubly labelled water. J Appl Physiol 1982;53:955–9.
20. Roberts SB, Coward WA, Schlingenseipen KH, Nohra V, Lucas A. Comparison of the doubly labeled water ($^2H_2{}^{18}O$) method with indirect calorimetry and a nutrient balance study for simultaneous determination of energy expenditure, water intake, and metabolizable energy intake in preterm infants. Am J Clin Nutr 1986;44:315–22.
21. Jones PJH, Winthrop AL, Schoeller DA, et al. Validation of doubly labelled water for assessing energy expenditure in infants. Pediatr Res 1987;21:242–6.
22. Schoeller DA, Webb P. Five-day comparison of the doubly labeled water method with respiratory gas exchange. Am J Clin Nutr 1984;40:153–8.
23. Klein PD, James WPT, Wong WW, et al. Calorimetric validation of the doubly labelled water method for determination of energy expenditure in man. Hum Nutr Clin Nutr 1984;38C:95–106.
24. Prentice AM, Coward WA, Davies HL, et al. Unexpectedly low levels of energy expenditure in healthy women. Lancet 1985;1: 1419–22.
25. Schoeller DA, Ravussin E, Schultz Y, Acheson KJ, Baertschi P, Jequier E. Energy expenditure by doubly labelled water: validation in humans and proposed calculation. Am J Physiol 1986;250: R823–30.
26. Arnould JPY. The energetics of lactation in polar bears (*Ursus maritimus* Phipps). MSc thesis. University of Saskatoon, Saskatchewan, Canada, 1990.
27. Westerterp KA, Kayser B, Brouns F, Herry JP, Saris WHM. Energy expenditure climbing Mt Everest. J Appl Physiol 1992;73:1815–9.
28. Westerterp KA, Meijer GAL, Janssen EME, Saris WNM, Ten Hoor F. Long-term effect of physical activity on energy balance and body composition. Br J Nutr 1992;68:21–30.
29. Black AE, Coward WA, Cole TJ, Prentice AM. Human energy expenditure in affluent society—an analysis of 574 doubly labelled water measurements. Eur J Clin Nutr 1996;50:72–92.
30. Schoeller DA. Measurement of energy expenditure in free-living humans by using doubly-labeled water. J Nutr 1988;118:1278–89.
31. Speakman JR. Doubly labelled water: theory and practice. London: Chapman and Hall, 1997.
32. Coward WA, Prentice MA, Murgatroyd PR, et al. Measurement of CO_2 and water production rates in man using ^2H, ^{18}O labelled H_2O: comparisons between calorimeter and isotope values. In: van Es AJH, ed. European Nutrition Report 5: human energy metabolism: physical activity and energy expenditure measurements in epidemiological research based upon direct and indirect calorimetry. The Hague: CIP-gegevens Koninklijke Bibliotheek, 1985.
33. Speakman JR. Principles, problems and a paradox with the measurement of energy expenditure of free-living subjects using doubly labelled water. Stats Med 1990;9:1365–80.
34. Speakman JR. Calculation of CO_2 production in doubly labelled water studies. J Theor Biol 1987;126:101–4.
35. Dykstra CR, Meyer MW, Karasov WH. Validation of the doubly labeled water method in bald eagles (*Haliaeetus leucocephalus*) and a comparison of two equations for the calculation of energy expenditure. Physiol Zool 1997;70:19–26.
36. Speakman JR, Nair KS, Goran MI. Revised equations for calculating CO_2 production from doubly labelled water in humans. Am J Physiol 1993;264:E912–7.
37. Matthews DE, Gilker CD. Impact of the H-2 and O-18 pool size determinations on the calculation of total energy expenditure. Obes Res 1995;3:S21–9.
38. Racette SB, Schoeller DA, Luke AH, Shay K, Hnilicka J, Kushner RF. Relative dilution spaces of ^2H- and ^{18}O-labeled water in humans. Am J Physiol 1994;30:E585–90.
39. Coward WA, Ritz P, Cole TJ. Revision of calculations of the doubly-labeled water method for measurement of energy expenditure in humans. Am J Physiol 1994;30:E805–7.
40. Westerterp KR, de Boer JO, Saris WHM, Schoffelen PFM, Ten Hoor F. Measurement of energy expenditure using doubly labeled water. Int J Sports Med 1984;5:74–5.
41. Schoeller DA, Kushner RF, Jones PJH. Validation of doubly labeled water for measuring energy expenditure during parenteral nutrition. Am J Clin Nutr 1986;44:291–8.
42. Westerterp KR, Brouns F, Saris WHM, Ten Hoor F. Comparison of doubly-labelled water with respirometry at low activity and high activity levels. J Appl Physiol 1988;65:53–6.
43. Jones PJ, Winthrop AL, Schoeller DA, et al. Evaluation of doubly labeled water for measuring energy expenditure during changing nutrition. Am J Clin Nutr 1988;47:799–804.
44. Seale JL, Miles C, Bodwell CE. Sensitivity of methods for calculating energy expenditure by use of doubly labeled water. J Appl Physiol 1989;66:644–53.
45. Seale JL, Rumpler WV, Conway JM, Miles CW. Comparison of doubly labeled water, intake balance, and direct- and indirect-calorimetery methods for measuring energy expenditure in adult men. Am J Clin Nutr 1990;52:66–71.
46. Ravussin E, Harper I, Rising R, Bogardus C. Energy expenditure by doubly labeled water: validation in lean and obese subjects. Am J Physiol 1991;261:E402–9.
47. Hoyt RW, Jones TE, Stein TP, et al. Doubly labeled water measurement of human energy expenditure during strenuous exercise. J Appl Physiol 1991;71:16–22.
48. Jones PJH, Leitch CA. Validation of doubly labeled water for measurement of caloric expenditure in collegiate swimmers. J Appl Physiol 1993;74:2909–14.
49. Seale JL, Conway JM, Canary JJ. Seven-day validation of doubly labeled water method using indirect room calorimetry. J Applied Physiol 1993;74:402–9.
50. Pullicino E, Coward WA, Elia M. Total energy expenditure in intravenously fed patients measured by the doubly labelled water technique. Metabolism 1993;42:58–64.
51. Speakman JR. How should we calculate CO_2 production in doubly labelled water studies of animals? Funct Ecol 1993;7:746–50.
52. Roberts SB, Dietz W, Sharp T, Dallal GE, Hill JO. Multiple laboratory comparison of the doubly labeled water technique. Obes Res 1995;3:S3–13.
53. Schoeller DA, Hnilicka JM. Reliability of the doubly-labeled water method for the measurement of total daily energy expenditure in free-living subjects. J Nutr 1996;126:S348–54.
54. Speakman JR. Estimation of precision in DLW studies using the two-point methodology. Obes Res 1995;3(suppl 1):31–9.

The American Journal of Clinical Nutrition

Downloaded from ajcn.nutrition.org by guest on January 31, 2013